

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: BIO MERIEUX  
(B) STREET: CHEMIN DE L'ORME  
(C) CITY: MARCY L'ETOILE  
(E) COUNTRY: FRANCE  
(F) POSTAL CODE: 69280

(ii) TITLE OF INVENTION: NUCLEIC MATERIAL OF THE  
ENDOGENOUS RETROVIRAL GENOMIC TYPE, ASSOCIATED WITH AN  
AUTOIMMUNE DISEASE AND/OR WITH PREGNANCY DISORDERS; USE  
AS MARKER

(iii) NUMBER OF SEQUENCES: 35

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0,  
Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1321 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: mRNA (as DNA)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CAACAATCGG GATATAAACC CAGGCATTGG AGCTGGCAAC AGCAGCCCCC CTTGGGTCC 60

CTTCCCTTG TATGGGAGCT GTTTTCATGC TATTTCACTC TATTAAATCT TGCAACTGCA 120

CTCTTCTGGT CCATGTTTCT TACGGCTCGA GCTGAGCTTT TGCTCACCGT CCACCACTGC 180

TGTTTGCCAC CACCGCAGAC CTGCCGCTGA CTCCCATCCC TCTGGATCCT GCAGGGTGTC 240

CGCTGTGCTC CTGATCCAGC GAAGCGCCCA TTGCCGCTCC CAATTGGGCT AAAGGCTTGC 300

CATTGTTCTT GCACGGCTAA GTGCCTGGGT TTGTTCTAAT TGAGCTGAAC ACTAGTCACT 360

GGGTTCCATG GTTCTCTTCT GTGACCCACG GCTTCTAATA GAACTATAAC ACTTACCACA 420

TGGCCCAAGA TTCCATTCTT TGGAAATCCGT GAGGCCAAGA ACTCCAGGTC AGAGAATACG 480

AAGCTTGCCA CCATCTTGGA AGCGGCCTGC TACCATCTTG GAAGTGGTTC ACCACCATCT 540

TGGGAGCTCT GTGAGCAAGG ACCCCCCGGT AACATTTTGG CAACCACGAA CGGACATCCA 600

AAGTGATGGG AAACGTTCCC CGCAAGACAA AAACGCCCCT AAGACGTATT CTGGAAAATT 660

GGGAACAATT TGACCCTCAG AACTAAGAA AGAAACGACT TATATTCTTC TGCAGTGCCG 720

CCTGGCACTC CTGAGGGAAG TATAAATTAT AACACCATCT TACAGCTAGA CCTCTTTTGT 780

AGAAAAGGCA AATGGAGTGA AGTGCCATAA GTACAACTT TCTTTTCATT AAGAGACAAC 840

TCACAATTAT GTAAAAGTG TGATTTATGC CCTACAGGAA GCCTTCAGAG TCTACCTCCC 900

TATCCCAGCA TCCCCGACTC CTCCCCCACT TAATAAGGAC CCCCCTTCAA CCCAAATGCT 960

CCAAAAGGAG ATAGACAAAA GGGTAAACAG TGAACCAAAG AGTGCCAATA TTCCCCAATT 1020

09445024 12169

ATGACCCCTC CAAGCAGTGG GAGGAAGAGA ATTCGGCCCA GCCAGAGTGC ATGTGCCTTT 1080  
 TTCTCTCCCA GACTTAAAGC AAATAAAAAC AGACTTAGGT AAATTCTCAG ATAACCCTGA 1140  
 TGGCTATATT GGTGTTTTAC AAGGGTTAGG ACAATTCTTT GATCTGACAT GGAGAGATAT 1200  
 ATATGTCACT GCTAAATCAG AACTAACCC CAAATGAGAG AAGTGCCACC ATAACGCAG 1260  
 CCTGAGAGTT TGGCGATCTC TGGTATCTCA GTCAGGTCAA TGATAGGATG ACAACAGAGG 1320  
 A 1321

(2) INFORMATION FOR SEQ ID NO: 2:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2938 base pairs  
 (B) TYPE: nucleotide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 10 (ii) MOLECULE TYPE: mRNA (as DNA)  
 (iii) HYPOTHETICAL: NO  
 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CAACGACGGA CATCCAAAGT GATGGGAAAC GTTCCCCGCA AGACAAAAAC GCCCCTAAGA 60  
 CGTATTCTGG AGAATTGGGA CCAATTTGAC CCTCAGACAC TAAGAAAGAA ACGACTTATA 120  
 TTCTTCTGCA GTGCCGCCTG GCACTCCTGA GGGAACTATA AATTATAACA CCATCTTACA 180  
 GCTAGACTTC TTTTGTAGAA AAGGCAAATG GAGTGAAGTG CCATAAGTAC AAACCTTCTT 240

TTCATTAAGA GACAACAC AATTATGTAA AAAGTGTGAT TTATGCCCTA CAGGAAGCCT 300

TCAGAGTCTA CCTCCCTATC CCAGCATCCC CGACTCCTTC CCCAACTAAT AAGGACCCCC 360

CTTCAACCCA AATGGTCCAA AAGGAGATAG ACAAAGGGT AAACAGTGAA CCAAAGAGTG 420

CCAATATTCC CCAATTATGA CCCCTCCCAA GCAGTGGGAG GAAGAGATTC GGCCCAGCCA 480

GAGTGCATGT GCTTTTCTT CTCCCAGACT TAAAGCAAAT AAAACAGAC TTAGGTAAAT 540

TCTCAGATAA TCCTGATGGC TATATTGATG TTTTACAAGG GTTAGGACAA TTCTTTGATC 600

TGACATGGAG AGATATAATG TCACTGCTAA ATCAGACACT AACCCCAAAT GAGAGAAGTG 660

CCACCATAAC TGCAGCCTGA GAGTTTGGCG ATCTCTGGTA TCTCAGTCAG GTCAATGATA 720

GGATGACAAC AGAGGAAAGA GATGATCCCC ACAGCCAGCA AGCAGTTCCC AGTCTASACC 780

CTCATTGGGG ACACAGAAAT CAGTAACATG GGAGATTGGT GCTGCAGACA TTTGCTAACT 840

TGTGTGCTAC AAGGACTAAG GAAACTACG AAGAAATCT ACGAATTACT CAATGATGTC 900

CACCATAACA CAGGGGAAGG GAAGAAATC CTAAGGAGAG CTAAGGGAGG 960

CATTGAGGAA GCGTGCCTCT CTGTCACCTG ACTCTTCTGA AGGCCAACTA ATCTTAAAGC 1020

GTAAGTTTAT CACTCAGTCA GCTGCAGACA TTAGAAAAA CTTCAAAGT CTGCCGTAGG 1080

CCCGGAGCAA AACTTAGAAA CCCTATTGAA CTTGGCAACY TCGGTTTTTT ATAATAGAGA 1140

TCAGGAGGAG CAGGCGGAAC AGGACAAACG GGATTAAAAA AAAGGCCACC GCTTTAGTCA 1200

TGACCCTCAG GCAAGTGGAC TTTGGAGGCT CTGGAAAAGG GAAAAGCTGG GCAAATTGAA 1260

TGCCTAATAG GGCTTGCTTC CAGTGGGTC TACAAGGACA CTTTAAAAA CATTGTCCAA 1320

669727-42094460

GTAGAAGTAA GCCGCCCCCTT CGTCCATGCC CTTATTTCA AGGGAATCAC TGGAAGGCCC 1380  
ACTGCCCCAG GGGACAAAGG TCTTTTGAGT CAGAAGCCAC TAACCAGATG ATCCAGCAGC 1440  
AGGACTGAGC GTGCCTGGGG CAAGCGCCAT CCCATGCCAT CACCCTCACA GAGCCCTGGG 1500  
TATGCTTGAC CATTGAGGGC CAGGAAGGTT GTCTCCTGGA CACTGGTGCG GTCTTCTTAG 1560  
TCTTACTCTT CTGTCCCGGA CAACTGTCCT CCAGATCTGT CACTATCTGA GGGGGTCCTA 1620  
AGACGGGCAG TCACTAGATA CTTCTCCAG CCACTAAGTT ATGACTGGGG AGCTTTATTC 1680  
TTTTCACATG CTTTTCTAAT TATGCTTGAA AGCCCCACTA CTTGTTAGG GAGAGACATT 1740  
CTAGCAAAAG CAGGGGCCAT TATACACCTG AACATAGGAG AAGGAACACC CGTTTGTGT 1800  
CCCCTGCTTG AGGAAGGAAT TAATCCTGAA GTCTGGGCAA CAGAAGGACA ATATGGACGA 1860  
GCAAAGAATG CCCGTCCTGT TCAAGTTAA CTAAAGGATT CCACTTCCTT TCCCTACCAA 1920  
AGGCAGTACC CCCTCAGACC CAAGGCCCAA CAAGGATTCC AAAAGATTGT TAAGGACTTA 1980  
AAAGCCCAAG GCTTAGTAAA ACCATGCATA ACTCCCTGCA GTAATTCCGT AGTGGATTGA 2040  
GGAGGCACAG AAACCCAGTC GACAGTGGAG GGTAGTGCA AGATCTCAGG ATTATCAATG 2100  
GAGGCCGTTG TCCTTTTATA CCCAGCTGTA CCTAGCCCTT ATACTGTGCT TTCCCAAATA 2160  
CCAGAGGAAG CAGAGTGGTT TACACTCCTG GACCTTAAGG ATGCCTTCTT CTGCATCCCT 2220  
GTACATCCTG ACTCTCAATT CTTGTTTGCC TTTGAAGATA CTTCAAACCC AACATCTCAA 2280  
CTCACCTGGA CTGTTTTACC CCAAGGGTTC AGGGATAGCC CCCATCTATT TGGCCAGGCA 2340  
TTAGCCCAAG ACTTGAGCCA ATCCTCATAC CTGGACACTT GTCCTTCGGT AGGTGGATGA 2400

TTTACTTTTG GCCGCCCAT CAGAAACCTT GTGCCATCAA GCCACCCAAG CGCTCTTCAA 2460

TTTCCTCGCT ACCTGTGGCT ACATGGTTTC CAAACCAAAG GCTCAACTCT GCTCACAGCA 2520

GGTTACTTAG GGCTAAAATT ATCCAAAGGC ACCAGGGCCC TCACTGAGGA ACACATCCAG 2580

CCTATACTGG CTTATCCTCA TCCCAAACC CTAAAGCAAC TAAGGGGATT CCTTGGCGTA 2640

ATAGGTTTCT GCCGAAAATG GATTCCCAGG TTTGGCGAAA TAGCCAGGTC ATTAAATACA 2700

CTAATTAAGG AACTCAGAA AGCCAATACC CATTAGTAA GATGGACAAC TGAAGTAGAA 2760

GTGGCTTTCC AGGCCCTAAC CCAAGCCCCA GTGTTAAGTT TGCCAACAGG GCAAGACTTT 2820

TCTTCATATG TCACAGAAAA AACAGGAATA GCTCTAGGAG TCCTTACACA GATCCGAGGG 2880

ATGAGCTTGC AACCTGTGGC GTACCTGACT AAGGAAATTG ATGTAGTGGC AAAGGGTT 2938

(2) INFORMATION FOR SEQ ID NO: 3:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1422 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: mRNA (as DNA)
- (iii) HYPOTHETICAL: NO
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TCAGGGATAG CCCCATCTA TTTGGCCAGG CATTAGCCCA AGACTTGAGT CAGTTATCAT 60

ACCTGGACAC TCTTGTCTT CAGTATGTGG ATGATTACT TTTAGCTGCC TGTTAGAAA 120

CCTTGTGCCA TCAAGCCACC CAAGCACTCT TAAATTTCTT CGCCACCTGT GGCTACAAGG 180

TTTCCAAAGA GAAGCTCAGC TCTGCTCACA GCAGGTAAA TACTTAGGAC TAAGATTATC 240

CAAAGGCACC AAGGCCCTCA GTGAGGAATG TATCCAGCCT ATACTGGCTT ATCCTCATCT 300

CAAAACCCTA AAGCAACTAA GAGAGTTCTT TGGCATAACA GGCTTCTGCC GAATATGGAT 360

TCCCCAGGTA TGGCAAATA GCCAGGCCAT TATATACAGT AATTAAGGAA ACTCAGAAAG 420

CCAATACCCA TTTAATAAGA TGGATACCTG AAGCCAAAGT GGCTTTCCAG GCCCCTAAAG 480

AAGGCCTTAA ACCCAAGTCC CAGTGTTAAG CTTGCCAACG GGGCAAGACT TTTCTTTATA 540

CATCACAGAA AAAACAGAA ACAGCTCTGG GAGTCCTTAC ACAGGTCCAA GGGACGAGCT 600

TGCAACCCAT GGCATACCTG AGTAAGGAAA CTGATGTAGT GGCAAAGGGT TGGCTTCATT 660

GTTTATGGGT AGTGCTGCCA GTAGCAGTTG TAGTATCTGA AGCAGTTAAA ATAATACAGG 720

GGAGAGATCT TACTGTGTGG ACATCTCATG AGGTGAACAG CATACTCACT GCTAAAGGAG 780

ACTTGTGGCT GTCAGACAAC CGTTTACTTA AATATCAGGC TCTATTACTT GAAAGGCCAG 840

TGCTGCAACT GTGCACTGT GCAACTCTTA ACCCAGTCNC ATTTCTTCCA GACAATGAAG 900

ATAGAATATA ACTGTCAACA AATAATTTCT CAAACCTATG CCACTCGAGG GGACCTTCTA 960

GAAGTTCCCT TGA CTGATCC TGACCTTCAA CTTGTATACT GATGGAAGTT CCTTTGTAGA 1020

AAAAGGACTT CAAAAGCGGG GTATGCAGTG GTCAGTGATA ATGGAATATT TGAAAGTATC 1080

CCCTCACTCC AGGAACTAGT GCTTAGCTGG CAGAACTAAT AGCCTTCATT GGGGCACTAG 1140

AATTAGGAGA AGGAAAAGC GTAAATATAT ATACAGACTC TGAGTATGCT CACCTAGTCN 1200

TCCATGCCCA TGAGGCAATA TGCAGAGAAA GGAATTCCT AACTCCGAG GGAACACCTA 1260  
TCACACATCA GGAAGCCATT AGGAGATTAT TACTGGCAGT ACAGAAACCT AAAGAGGTGG 1320  
AAGTCTTACA CTGCTGGGGT CATCAGAAAG GAAAGAAAAG GGAAATAGAA GGAATTGCC 1380  
AAGCAGATAT TGAAGCAAAA AGAGCTGCAA GGCAGGACCC TC 1422

(2) INFORMATION FOR SEQ ID NO: 4:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2006 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

- (ii) MOLECULE TYPE: mRNA (as DNA)  
(iii) HYPOTHETICAL: NO

15

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATGCAGTGGT CAGTGATAAT GGAATACTTG AAAGTAATCC CCTCACTCCA GGAAGTAGTG 60  
CTCAGCTAGC AGAACTAATA GCCCTCACTT GGGCACTAGA ATTAGGAGAA GAAAAAGGG 120  
CAAATATATA TACAGACTCT AAATATGCTT ACCTAGTCCT CCATGCCCAT GCAGCAATAT 180  
GGAAAGAAAAG GGAATTCCTA ACTTCTGAGA GAACACCTAT CAAACATCAG GAAGCCATTA 240  
GGAAATTATT ATTGGCTGTA CAGAAACCTA AAGAGGTGGC AGTCTTACAC TGCCGGGGTC 300  
ATCANAAAGG AAAGGAAAGG GAAATACTT TTGCETGCAA CTATCCAATG GAAATTACTT 360



AAAACCCTTC ATCAAACCTT TCACTTAGGC ATCGATAGCA CCCATCAAAT GGCCAAATCA 420  
 TTATTTACTG GACCAGGCCT TTTCAAACT ATCAAGCAAA TATTCAGGGC CTGTGAATTG 480  
 TGCCAAAAAA ATAATCCCCT GCCTCATCGC CAAGCTCCTT CAGGAAAACA AAAAACAGGC 540  
 CATTACCCTG AAAAAAAGTG GCAACTGATT TTACCCACAA GCCCAAACCT CAGGGATTTC 600  
 AGTATCTACT AGTCTGGGTA AATACTTTCA CGGGTTGGGC AAAGGCCTTC CCCTGTAGGA 660  
 CAGAAAAGGC CCAAGAGGTA ATAAAGGCAC TAGTTCATGA AATAATTCCC AGATTCCGAC 720  
 TTCCCGAGG CTTACAGAGT GACAATAGCC CTGCTTTCCA GGCCACAGTA ACCCAGGGAG 780  
 TATCCAGGC GTTAGGTATA CGATATCACT TACACTGCGC CTGAAGGCCA CAGTCCTCAG 840  
 GGAAGGTCGA GAAAATGAAT GAAATACTCA AAGGACATCT AAAAAAGCAA ACCCAGGAAA 900  
 CCCACCTCAC ATGGCCTGCT CTGTTGCCTA TAGCCTTAAA AAGAATCTGC AACTTTCCCC 960  
 AAAAGCAGG ACTTAGCCCA TACGAAATGC TGTATGGAAG GCCCTTCATA ACCAATGACC 1020  
 TTGTGCTTGA CCCAAGACAG CCAACTTAGT TGCAGACATC ACCTCCTTAG CCAAATATCA 1080  
 ACAAGTTCTT AAAACATTAC AAGGAACCTA TCCCTGAGAA GAGGGAAAAG AACTATTCCA 1140  
 CCCTTGAGAC ATGGTATTAG TCAAGTCCCT TCTCTCTAAT TCCCCATCCC TAGATACATC 1200  
 CTGGGAAGGA CCCTACCCAG TCATTTTATT TACCCCAACT GCGGTAAAG TGGCTGGAGT 1260  
 GGTCTTGGAT ACATCACACT TGAGTCAAAT CCTGGATACT GCCAAAGGAA CCTGAAAATC 1320  
 CAGGAGACAA CGCTAGCTAT TCCTGTGAAC CTCTAGAGGA TTTGCGCCTG CTCTTCAAAC 1380  
 AACAACCAGG AGGAAAGTAA CTAAAATCAT AAATCCCCCA TGGCCCTCCC TTATCATATT 1440

09446024.121699

TTTCTCTTTA CTGTTCTTTT ACCCTCTTTC ACTCTCACTG CACCCCCTCC ATGCCGCTGT 1500  
 ATGACCAGTA GCTCCCCCTTA CCAAGAGTTT CTATGGAGAA TGCAGCGTCC CGGAAATATT 1560  
 GATGCCCCAT CGTATAGGAG TCTTTCTAAG GGAACCCCCA CCTTCACTGC CCACACCCAT 1620  
 ATGCCCCGCA ACTGCTATCA CTCTGCCACT CTTTGCATGC ATGCAAATAC TCATTATTGG 1680  
 ACAGGAAAAA TGATTAATCC TAGTTGTCCT GGAGGACTTG GAGTCACTGT CTGTTGGACT 1740  
 TACTTCACCC AAAGTGGTAT GTCTGATGGG GGTGGAGTTC AAGATCAGGC AAGAGAAAAA 1800  
 CATGTAAAAG AAGTAATCTC CCAACTCACC CGGGTACATG GCACCTCTAG CCCTACAAAG 1860  
 GACTAGATCT CTCAAAACTA CATGAAACCC TCCGTACCCA TACTCGCCTG GTAAGCCTAT 1920  
 TTAATACCAC CCTCACTGGG CTCCATGAGG TCTCGGCCCA AAACCCTACT AACTGTTGGA 1980  
 TATGCCTCCC CCTGAACTTC AAGCCA 2006

(2) INFORMATION FOR SEQ ID NO: 5:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1948 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: mRNA (as DNA)

(iii) HYPOTHETICAL: NO

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ACTGCACTCT TCTGGTCCAT GTTCTTACG GCTCGAGCTG AGCTTTTGCT CACCGTCCAC 60

CACTGCTGTT TGCCACCACC GCANACCTGC CGCTGACTCC CATCCCTCTG GATCCTGCAG 120

GGTGTCGCT GTGCTCCTGA TCCAGCGAGG CGCCCATTGC CGCTCCCAAT TGGGCTAAAG 180

GCTTGCCATT GTNCCTGCAC GGCTAAGTGC CTGGGTTTGT TCTAATTGAG CTGAACACTA 240

NTCACTGGGT TCCATGGTTC TCTTCTGTGA CCCACGGCTT CTAATAGAAC TATAACACTT 300

ACCACATGGC CCAAGATTCC ATTCCTTGGA ATCCGTGAGG GCAAGAACTC CAGGTCAGAG 360

AATACGAGGC TTGCCACCAT CTTGGAAGCG GCCTGCTACC ATCTTGGAAG TGGTTCACCA 420

CCATCTTGGG AGCTCTGTGA GCAAGGACCC CCCGGTAACA TTTTGGAAC CACGAACGGA 480

CATCCAAAGT GATACATCCT GGAAGGACC CTACCCAGTC ATTTTATCTA CCCCAACTGC 540

GGTTAAAGTG GCTGGAGTGG AGTCTTGAT ACATCACACT TGAGTCAAAT CCTGGATACT 600

GCCAAAGGAA CCTGAAAATC CAGGAGACAA CGCTAGCTAT TCCTGTGAAC CTCTAGAGGA 660

TTTGCGCCTG CTCTTCAAAC AACAAACCAGG AGGAAAGTAA CTAAAATCAT AAATCCCCAT 720

GGCCCTCCCT TATCATATTT TTCTCTTTAC TGTTGTTTCA CCCTCTTTCA CTCTCACTGC 780

ACCCCTCCA TGCCGCTGTA TGACCAGTAG CTCCCCTTAC CAAGAGTTTC TATGGAGAAT 840

GCAGCGTCCC GGAATATTG ATGCCCCATC GTATAGGAGT CTTTGTAAGG GAACCCCCAC 900

CTTCACTGCC CACACCCATA TGCCCCGCAA CTGCTATCAC TCTGCCACTC TTTGCATGCA 960

TGCAAATACT CATTATTGGA CAGGAAAAAT GATTAATCCT AGTTGTCCTG GAGGACTTGG 1020

AGTCACTGTC TGTTGGACTT ACTTCACCCA AACTGGTATG TCTGATGGGG GTGGAGTTCA 1080

AGATCAGGCA AGAGAAAAAC ATGTAAAAGA AGTAATCTCC CAACTCACCC GGGTACATGC 1140

CACCTCTAGC CCCTACAAAG GACTAGATCT CTCAAAATA CATGAAACCC TCCGTACCCA 1200  
 TACTCGCCTG GTAAGCCTAT TTAATACCAC CCTCACTGGG CTCCATGAGG TCTCGGCCCA 1260  
 AAACCCTACT AACTGTTGGA TATGCCTCCC CCTGAACTTC AGGCCATATG TTTCAATCCC 1320  
 TGTACCTGAA CAATGGAACA ACTTCAGCAC AGAAATAAAC ACCACTTCCG TTTTAGTAGG 1380  
 ACCTCTTGTT TCCAATCTGG AAATAACCCA TACCTCAAAC CTCACCTGTG TAAAATTTAG 1440  
 CAATACTACA TACACAACCA ACTCCCAATG CATCAGGTGG GTAACCTCTC CCACACAAAT 1500  
 AGTCTGCCTA CCCTCAGGAA TATTTTTTGT CTGTGGTACC TCAGCCTATC GTTGTTTGAA 1560  
 TGGCTCTTCA GAATCTATGT GCTTCCTCTC ATTCTTAGTG CCCCCTATGG CCATCTACAC 1620  
 TGAACAAGAT TTATACAGTT ATGTCATATC TAAGCCCCGC AACAAAAGAG TACCCATTCT 1680  
 TCCTTTTGTT ATAGGAGCAG GAGTGCTAGG TGCACTAGGT ACTGGCATTG GCGGTATCAC 1740  
 AACCTCTACT CAGTTCTACT ACAAATATC TCAAGAACTA AATGGGGACA TGGAACGGGT 1800  
 CGCCGACTCC CTGGTCACCT TGCAAGATCA ACTTAACTCC CTAGCAGCAG TAGTCCTTCA 1860  
 AAATCGAAGA GCTTTAGACT TGCTAACCGC TGAAGAGGGG GGAACCTGTT TATTTTTAGG 1920  
 GGAAGAATGC TGTTATTATG TTAATCAA 1948

(2) INFORMATION FOR SEQ ID NO: 6:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1136 base pairs  
 (B) TYPE: nucleotide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: mRNA (as DNA)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCATGGCCAT CTACACTGAA CAAGATTTAT ACAGTTATGT CATATCTAAG CCCC GCAACA 60

AAAGAGTACC CATTCTTCCT TTTGTTATAG GAGCAGGAGT GCTAGGTGCA CTAGGTA CTG 120

GCATTGGCGG TATCACAACC TCTACTCAGT TCTACTACAA ACTATCTCAA GAACTAAATG 180

GGGACATGGA ACGGGTCGCC GACTCCCTGG TCACCTTGCA AGATCAACTT AACTCCCTAG 240

CAGCAGTAGT CCTTCAAAAT CGAAGAGCTT TAGACTCGCT AACCGCTGAA AGAGGGGGAA 300

CCTGTTTATT TTTAGGGGAA GAATGCTGTT ATTATGTTAA TCAATCCGGA ATCGTCACTG 360

AGAAAGTTAA AGAAATTCGA GATCGAATAC AACGTAGAGC AGAAGAGCTT CGAAACACTG 420

GACCCTGGGG CCTCCTCAGC CAATGGATGC CCTGGATTCT CCCCTTCTTA GGACCTCTAG 480

CAGCTATAAT ATTGCTACTC CTCTTTGGAC CCTGTATCTT TAACCTCCTT GTTAACCTTG 540

TCTCTTCAG AATCGAAGCT GTAAAACTAC AAATGGAGCC CAAGATGCAG TCCAAGACTA 600

AGATCTACCG CAGACCCCTG GACCGGCCTG CTAGCCCACG ATCTGATGTT AATGACATCA 660

AAGGCACCCC TCCTGAGGAA ATCTCAGCTG CACAACCTCT ACTACGCCCC AATTCAGCAG 720

GAAGCAGTTA GAGCGGTCGT CGGCCAACCT CCCCAACAGC ACTTAGGTTT TCCTGTTGAG 780

ATGGGGGACT GAGAGACAGG ACTAGCTGGA TTTCCTAGGC TGA CTAAGAA TCCCTAAGCC 840

TAGCTGGGAA GGTGACCACA TCCACCTTTA AACACGGGGC TTGCAACTTA GTTCACACCT 900

GACCAATCAG AGAGCTCACT AAAATGCTAA TTAGGCAAAG ACAGGAGGTA AAGAAATAGC 960

CAATCATCTA TTGCATGAGA GCACAGCAGG AGGGACAATG ATCGGGATAT AAACCCAAGT 1020

CTTCGAGCCG GCAACGGCAA CCCCCTTTGG GTCCCTCCC TTTGTATGGG AGCTCTGTTT 1080

TCATGCTATT TCACTCTATT AAATCTTGCA GCTGCCAAAA AAAAAAAAAA AAAAAA 1136

5

09446024.12169

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 2782 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: mRNA (as DNA)

10

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

669121" 1209460  
15 ATGGGAGCTG TTTTCATGCT ATTTCACTCT ATTAAATCTT GCAACTGCAC TCTTCTGGTC 60  
CATGTTTCTT ACGGCTCGAG CTGAGCTTTT GCTCACCGTC CACCACTGCT GTTTGCCACC 120  
ACCGCAGACC TGCCGCTGAC TCCCATCCCT CTGGATCCTG CAGGGTGTCC GCTGTGCTCC 180  
TGATCCAGCG AAGCGCCCAT TGCCGCTCCC AATTGGGCTA AAGGCTTGCC ATTGTTCTTG 240  
CACGGCTAAG TGCCTGGGTT TGTCTAATT GAGCTGAACA CTACTCACTG GGTTCATGG 300  
TTCTCTTCTG TGACCCACGG CTTCTAATAG AACTATAACA CTTACCACAT GGCCCAAGAT 360

TCCATTCCCTT GGAATCCGTG AGGCCAACGA ACTCCAGGTC AGAGAATACG AAGCTTGCCA 420

CCATCTTGGA AGCGGCCTGC TACCATCTTG GAAGTGGTTC ACCACCATCT TGGGAGCTCT 480

GTGAGCAAGG ACCCCCCGGT GACATTTTGG CGACCACCAA CGGACATCCC AAGTGATACA 540

TCCTGGGAAG GACCCTACCC AGTCATTTTA TCTACCCCAA CTGCGGTAA AGTGGCTGGA 600

GTGGAGTCTT GGATACATCA CACTTGAGTC AAATCCTGGA TACTGCCAAA GGAACCTGAA 660

AATCCAGGAG ACAACGCTAG CTATTCCTGT GAACCTCTAG AGGATTTGCG CCTGCTCTTC 720

AAACAACAAC CAGGAGGAAA GTAACATAAA TCATAAATCC CCATGGGCCT CCCTTATCAT 780

ATTTTTCTCT GTAGTGTTCT TTCACCCTGT TTCACTCTCA CTGCACCCCC TCCATGCCGC 840

TGTATGACCA GTAGCTCCCC TCACCCAGAG TTTCTATGGA GAATGCAGCG TCCCGGAAAT 900

ATTGATGCCC CATCGTATAG GAGTCTTTCT AAGGGAACCC CCACCTTAC TGCCACACC 960

CATATGCCCC GCAACTGCTA TCACTCTGCC ACTCTTTGCA TGCATGCAA TACTCATTAT 1020

TGGACAGGAA AAATGATTAA TCCTAGTTGT CCTGGAGGAC TTGGAGTCAC TGTCTGTTGG 1080

ACTTACTTCA CCCAACTGG TATGTCTGAT GGGGGTGGAG TTCAAGATCA GGCAAGAGAA 1140

AAACATGTAA AAGAAGTAAT CTCCCAACTC ACCGGGGTAC ATGGCACCTC TAGCCCCTAC 1200

AAAGGACTAG ATCTCTCAA ACTACATGAA ACCCTCCGTA CCCATACTCG CCTGGTAAGC 1260

CTATTTAATA CCACCCTCAC TGGGCTCCAT GAGGTCTCGG CCCAAAACCC TACTAACTGT 1320

TGGATATGCC TCCCCCTGAA CTTCAGGCCA TATGTTTCAA TCCCTGTACC TGAACAATGG 1380

AACAACCTCA GCACAGAAAT AACACCACT TCCGTTTTAG TAGGACCTCT TGTTTCCAAT 1440

09445024 121699

GTGGAAATAA CCCATACCTC AAACCTCACC TGTGTAAAAT TTAGCAATAC TACATACACA 1500  
ACCAACTCCC AATGCATCAG GTGGGTAACCT CCTCCACAC AAATAGTCTG CCTACCCTCA 1560  
GGAATATTTT TTGTCTGTGG TACCTCAGCC TATCGTTGTT TGAATGGCTC TTCAGAATCT 1620  
ATGTGCTTCC TCTCATTTCTT AGTGCCCCCT ATGACCATCT ACACTGAACA AGATTTATAC 1680  
AGTTATGTCA TATCTAAGCC CCGCAACAAA AGAGTACCCA TTCTTCCTTT TGTTATAGGA 1740  
GCAGGAGTGC TAGGTGCACT AGGTACTGGC ATTGGCGGTA TCACAACCTC TACTCAGTTC 1800  
TACTACAAAC TATCTCAAGA ACTAAATGGG GACATGGAAC GGGTCGCCGA CTCCCTGGTC 1860  
ACCTTGCAAG ATCAACTTAA CTCCCTAGCA GCAGTAGTCC TTCGAAATCG AAGAGCTTTA 1920  
GACTTGCTAA CCGCTGAGAG AGGGGGAACC TGTTATTTT TAGGGGAAGA ATGCTGTTAT 1980  
TATGTTAATC AATCCGGAAT CGTCACTGAG AAAGTTGAAG AAATTCCAGA TCGAATACAA 2040  
CGTATAGCAG AGGAGCTTCG AAACACTGGA CCCTGGGGCC TCCTCAGCCG ATGGATGCCC 2100  
TGGATTCTCC CCTTCTTAGG ACCTCTAGCA GCTATAATAT TGCTACTCCT CTTTGGACCC 2160  
TGATCTTTG ACCTCCTTGT TAACTTTGTC TCTCCAGAA TCGAAGCTGT GAAACTACAA 2220  
ATGGAGCCCA AGATGCAGTC CAAGACTAAG ATCTACCGCA GACCCCTGGA CCGGCCTGCT 2280  
AGCCCACGAT CTGATGTTAA TGACATCAAA GGCACCCCTC CTGAGGAAAT CTCAGCTGCA 2340  
CAACCTCTAC TACGCCCCAA TTCAGCAGGA AGCAGTTAGA GCGGTGGTCG GCCAACCTCC 2400  
CCAACAGCAC TTAGGTTTTT CTGTTGAGAT GGGGGACTGA GAGACAGGAC TAGCTGGATT 2460  
TCCTAGGCTG ACTAAGAATC CTTAAGCCTA GGTGGGAAGG TGACCACATC CACCTTTAAA 2520



CACGGGGCTT GCAACTTAGC TCACACCTGA CCAATCAGAG AGCTCACTAA AATGCTAATT 2580  
AGGCAAAGAC AGGAGGTAAA GAAATAGCCA ATCATTATT GCCTGAGAGC ACAGCAGGAG 2640  
GGACAATGAT CGGGATATAA ACCCAAGTTT TCGAGCCGGC AACGGCAACC CCCTTTGGGT 2700  
CCCCTCCCTT TGTATGGGAG CTCTGTTTTC ATGCTATTTC ACTCTATTAA ATCTTGCAAC 2760  
TGCAAAAAAA AAAAAAAAAA AA 2782

(2) INFORMATION FOR SEQ ID NO: 8:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 666 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: mRNA (as DNA)  
(iii) HYPOTHETICAL: NO

- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TGTCCGCTGT GTCCTGATC CAGCGAGGCG CCCATTGCCG CTCCCAATTG GGCTAAAGGC 60  
TTGCCATTGT TCCTGCACGG CTAAGTGCCT GGGTTTGTTT TAATTGAGCT GAACACTANT 120  
CACTGGGTTC CATGGTTCTC TTCTGTGACC CACGGCTTCT AATATAACTA TAACACTTAC 180  
CACATGGCCC AAGATTCCAT TCCTTGGAAT CCGTGAGGCC AAGAACTCCA GGTCAGAGAA 240  
TACGAGGCTT GCCACCATCT TGGAAGCGGC CTGCTACCAT CTTGGAAGTG GTTACCACC 300

ATCTTGGGAG CTCTGTGAGC AAGGACCCCC CGGTAACATT TTGGCAACCA CGAACGGACA 360  
TCCAAAGTGA ATCGAAGCTG TAAAACTACA AATGGAGCCC AAGATGCAGT CCAAGACTAA 420  
GATCTACCGC AGACCCCTGG ACCGGCCTGC TAGCCCACGA TCTGATGTTA ATGACATCAA 480  
AGGCACCCCT CCTGAGGAAA TCTCAGCTGC ACAACCTCTA CTACGCCCCA ATTCAGCAGG 540  
AAGCAGTTAG AGCGGTCGTC GGCCAACCTC CCCAACAGCA CTTAGGTTTT CTGTTGAGA 600  
TGGGGGACTG AGAGACAGGA CTAGCTGGAT TTCCTAGGCT GACTAAGAAT CCCTAAGCCT 660  
AGCTGG 666

(2) INFORMATION FOR SEQ ID NO: 9:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3372 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: mRNA (as DNA)
- (iii) HYPOTHETICAL: NO
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GACTTCCCAA ATACCAGAGG AAGCAGAGTG GTTTACAGTC CTGGACCTTC AGGATGCCTT 60  
CTTCTGCATC CCTGTACATC CTGACTCTCA ATTCTTGTTT GCCTTTGAAG ATACTTCAA 120  
CCCAGCATCT CAACTCACCT GGACTATTTT ACCCCAAGGG TTCAGGGATA GTCCCCATCT 180  
ATTTGGCCAG GCATTAGCCC AAGACTTGAG CCAATCCTCA TACCTGGACA CTTGTCCTTC 240

GGTAGGTGGA TGATTTACTT TTGGCCGCCC ATTACAGAAAC CTTGTGCCAT CAAGCCACCC 300

AAGCGCTCTT CAATTTCTC GCTACCTGTG GCTACATGGT TTCCAAACCA AAGGCTCAAC 360

TCTGCTCACA GCAGGTACT TAGGGCTAAA ATTATCCAAA GGCACCAGGG CCCTCAGTGA 420

GGAACACATC CAGCCTATAC TGGCTTATCC TCATCCCAA ACCCTAAAGC AACTAAGGGG 480

ATTCCTTGGC GTAATAGGTT TCTGCCGAAA ATGGATTCCC AGGTATGGCG AAATAGCCAG 540

GTCATTAAAT AACTAATTA AGGAACTCA GAAAGCCAAT ACCCATTTAG TAAGATGGAC 600

AACTGAAGTA GAAGTGGCTT TCCAGGCCCT AACCCAAGCC CCAGTGTTAA GTTTGCCAAC 660

AGGGCAAGAC TTTTGTTTAT ATGTCACAGA AAAAACAGGA ATAGCTCTAG GAGTCCTTAC 720

ACAGATCCGA GCGATGAGCT TGCAACCTGT GGCACACCTG ACTAAGGAAA TTGATGTACT 780

GGCAAAGGGT TGACCTCATT GTTTACGGGT AGTGGTGGCA GTAGCAGTCT TAGTATCTGA 840

AGCAGTTAAA ATAATACAGG GAAGAGATCT TACTGTGTGG ACATCTCATG ATGTGAATGG 900

CATACTCACT GCTAAAGGAG ACTTGTGGCT GTCAGACAAC TGTTTACTTA AATGTCAGGC 960

TCTATTACTT GAAGGGCCAG TGCTCGGACT GTGCACTTGT GCAACTCTTA ACCCAGCCAC 1020

ATTTCTTCCA GACAATGAAG AAAAGATAAA ACATAACTGT CAACAAGTAA TTTCTCAAAC 1080

CTATGCCACT CGAGGGGACC TTTTAGAGGT TCCTTTGACT GATCCCGACC TCAACTTGTA 1140

TACTGATGGA AGTTCCTTTG TAGAAAAAGG ACTTCGAAAA GTGGGGTATG CAGTGGTCAG 1200

TGATAATGGA ATACTTGAAA GTAATCCCCT CACTCCAGGA ACTAGTGCTC AGCTAGCAGA 1260

ACTAATAGCC CTCCTTGGG CACTAGAATT AGGAGAAGAA AAAAGGGCAA ATATAATACA 1320

GA CTCTAAAT ATGCTTACCT AGTCCTCCAT GCCCATGCAG CAATATGGAA AGAAAGGGAA 1380  
TTCCTAACTT CTGAGAGAAC ACCTATCAAA CATCAGGAAG CCATTAGGAA ATTATTATTG 1440  
GCTGTACAGA AACCTAGAGA GGTGGCAGTC TTACTGTC GGGGTCATCA CAAAGGAAAG 1500  
GAAAGGGAAA TACAAGAGAA CTGCCAAGCA TATATTGAAG CCAAAAGAGC TGCAAGGCAG 1560  
GACCCTCCAT TAGAAATGCT TATTAACTT CCCTTAGTAT AGGGTAATCC CTTCCGGGAA 1620  
ACCAAGCCCC AGTACTCAGC AGGAGAAACA GAATGGGGAA CCTCACGAGC CAGTTTTCTC 1680  
CCCTCGGGAC GGTTAGCCAC TGAAGAAGG AAAATACTTT TGCCTGCAAC TATCCAATGG 1740  
AAATTACTTA AAACCCTTCA TCAAACCTT CACTTAGGCA TCGATAGCAC CCATCAGATG 1800  
GCCAATCAT TATTTACTGG ACCAGGCCTT TTCAAACTA TCAAGCAGAT AGTCAGGGCC 1860  
TGTGAAGTGT GCCAGAGAAA TAATCCCCTG CCTTATCGCC AAGCTCCTTC AGGAGAACAA 1920  
AGAACAGGCC ATTACCCTGG AGAAGACTGG CAACTGATTT TACCCACAAG CCCAACCTC 1980  
AGGGATTCA GSTATCTACTA GTCTGGGTAG ATACTTTCAC GGGTTGGGCA GAGGCCTTCC 2040  
CCTGTAGGAC AGAAAAGGCC CAAGAGGTAA TAAAGGCACT AGTTCATGAA ATAATTCCCA 2100  
GATTCGGACT TCCCCGAGGC TTACAGAGTG ACAATAGCCC TGCTTTCAG GCCACAGTAA 2160  
CCCAGGGAGT ATCCCAGGCG TTAGGTATAC GATATCACTT AACTGCGCC TGAAGGCCAC 2220  
AGTCCTCAGG GAAGGTCGAG AAAATGAATG AAACACTCAA AGGACATCTA AAAAGCAAA 2280  
CCCAGGAAAC CCACCTCACA TGGCCTGTTC TGTTCCTAT AGCCTTAAA AGAATCTGCA 2340  
ACTTTCCCA AAAAGCAGGA CTTAGCCCAT ACGAAATGCT GTATGGAAG CCCTTCATAA 2400

CCAATGACCT TGTGCTTGAC CCAAGACAGC CAACTTAGTT GCAGACATCA CCTCCTTAGC 2460

CAAATATCAA CAAGTTCTTA AAACATTACA AGGAACCTAT CCCTGAGAAG AGGAAAAGAA 2520

TATTCCACCC AAGTGACATG GTATTAGTCA AGTCCCTTCC CTCTAATTCC CCATCCCTAG 2580

ATACATCCTG GGAAGGACCC TACCCAGTCA TTTTATCTAC CCCAACTGCG GTTAAAGTGG 2640

CTGGAGTGGA GTCTTGATA CATCACACTT GAGTCAAATC CTGGATACTG CCAAAGGAAC 2700

CTGAAAATCC AGGAGACAAC GCTAGCTATT CCTGTGAACC TCTAGAGGAT TTGCGCCTGC 2760

TCTTCAAACA ACAACCAGGA GGAAAAATCG AAGCTGTAAA ACTACAAATG GAGCCCAAGA 2820

TGCAGTCCAA GACTAAGATC TACCGCAGAC CCCTGGACCG GCCTGTTAGC CCACGATCTG 2880

ATGTTAATGA CATCAAAGGC ACCCCTCCTG AGGAAATCTC AGCTGCACAA CCTCTACTAC 2940

GCCCCAATTC AGCAGGAAGC AGTTAGAGCG GTCGTGGGCC AACCTCCCCA ACAGCACTTA 3000

GGTTTTCTG TTGAGATGGG GGAAGTGAAG ACAGGACTAG CTGGATTTCC TAGGCTGATT 3060

AAGAATCCCT AAGCCTAGCT GGAAGGTGA CCACATCCAC CTTTAAACAC GGGGCTTGCA 3120

ACTTAGCTCA CACCTGACCA ATCAGAGAGC TCACTAAAAT GCTAATTAGG CAAAGACAGG 3180

AGGTAAAGAA ATAGCCAATC ATTTATTGCC TGAGAGCACA GCAGGAGGGA CAATGATCGG 3240

GATATAAACC CAAGTTTTCG AGCCGGCAAC GGCAACCCCC TTTGGGTCCC CTCCCTTTGT 3300

ATGGGAGCTC TGTTTTCATG CTATTTCACT CTATTAAATC TTGCAACTGC AAAAAAAAAA 3360

AAAAAAAAAA AA 3372

(2) INFORMATION FOR SEQ ID NO: 10:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2372 base pairs
  - (B) TYPE: nucleotide
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

REPLACEMENT SHEET (RULE 26)

0944604.12169

(ii) MOLECULE TYPE: mRNA (as DNA)

(iii) HYPOTHETICAL: NO

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ACTGCACTCT TCTGGTCCAT GTTCTTACG GCTCGAGCTG AGCTTTTGCT CACCGTCCAC 60  
CACTGCTGTT TGCCACCACC GCAGACCTGC CGCTGACTCC CATCCCTCTG GATCCTGCAG 120  
GGTGTCCGCT GTGCTCCTGA TCCAGCGAGG CGCCCATTCG CGCTCCCAAT TGGGCTAAAG 180  
GCTTGCCATT GTTCCTGCAC GGCTAAGTGC CTGGGTTTGT TCTAATTGAG CTGAACACTA 240  
ATCACTGGGT TCCATGGTTC TCTTCTGTGA CCCACGGCTT CTAATAGAAC TATAACACTT 300  
ACCACATGGC CCAAGATTCC ATTCCTTGGA ATCCGTGAGG CCAAGAAGTC CAGGTCAGAG 360  
AATACGAGGC TTGCCACCAT CTGGAAGCG GCCTGCTACC GTCTTGGAAG TGGTTCACCA 420  
CCATCTTGGG AGCTCTGTGA GCAAGGACCC CCCGGTAACA TTTTGGCAAC CAACGACGGA 480  
CATCCAAAGT GATGGGAAAC GTTCCCCGCA AGACAAAAC GCCCCTAAGA CGTATTCTGG 540  
AGAATTGGGA CCAATTTGAC CCTCAGACAC TAAGAAAGAA ACGACTTATA TTCTTCTGCA 600  
GTGCCGCCTG GCACTCCTGA GGAAGTATA AATTATAACA CCATCTTACA GCTAGACCTC 660

0944604.2169  
669727-12094460

TTTTGTAGAA AAGGCAAATG GAGTGAAGTG CCATAAGTAC AAACTTTCTT TTCATTAAGA 720

GACAACTCAC AATTATGTAA AAAGTGTGAT TTATGCCCTA CAGGAAGCCT TCAGAGTCTA 780

CCTCCCTATC CCAGCATCCC CGACTCCTTC CCCAACTAAT AAGGACCCCC CTTCAACCCA 840

AATGGTCCAA AAGGAGATAG ACAAAGGGT AAACAGTGAA CCAAAGAGTG CCAATATTCC 900

CCAATTATGA CCCCTCCAAG CAGTGGGAGG AAGAGAATTC GGCCAGCCA GAGTGCATGT 960

GCCTTTTTCT CTCCCAGACT TAAAGCAAAT AAAACAGAC TTAGGTAAAT TCTCAGATAA 1020

CCCTGATGGC TATATTGATC TTTTACAAGG GTTAGGACAA TTCTTTGATC TGACATGGAG 1080

AGATATAATG TCACTGCTAA ATCAGACACT AACCCCAAAT GAGAGAAGTG CCACCATAAC 1140

TGCAGCCTGA GGGTTTGGCG TCTCTGGTAT CTCAGTCAGG TCAATGGATA NGGATGACAA 1200

CAGAAGGAAA GANAATGATT CCCCACAGGC CAGCAGGCAG TTCCCAGTCT AGACCCTCAT 1260

TGGGACACAG AATCAGAACA TGGAGATTGG TGCTGCAGAC ATTTGCTAAC TTGTGTGCTA 1320

GAAGGACTAA GGAAACTAG GAAGAAGTCT ATGAATTACT CAATGATGTC CACCATAACA 1380

CAGGGAAGGG AAGAAAATCC TACTGCCTTT CTGGAGAGAC TAAGGGAGGC ATTGAGCAAG 1440

CGTGCCTCTC TGTACCTGA CTCTTCTGAA GGCCAACTAA TCTTAAAGCG TAAGTTTATC 1500

ACTCAGTCAG CTGCAGACAT TAGAAAAAAC TTCAAAAGTC TGCCGTAGGC CCGGAGCAA 1560

ACTTAGAAAC CCTATTGAAC TTGGCAACCT CGGTTTTTTA TAATAGAGAT CAGGAGGAGC 1620

AGGCGGAACA GGACAAACGG GATTAAAAAA AAGGCCACCG CTTTAGTCAT GACCCTCAGG 1680

CAAGTGGACT TTGGAGGCTC TGGAAAAGGG AAAAGCTGGG CAAATTGAAT GCCTAATAGG 1740

66922"42094460

GCTTGCTTCC AGTGCGGTCT ACAAGGACAC TTAAAAAAG ATTGTCCAAG TAGAAGTAAG 1800  
CCGCCCCCTTC GTCCATGCCC CTTATTTCAA GGAATCACT GGAAGGCCCA CTGCCCCAGG 1860  
GGACAAAGGT CTTTTGAGTC AGAAGCCACT AACCAGATGA TCCAGCAGCA GGACTGAGGG 1920  
TGCCTGGGGC AAGCGCCATC CCATGCCATC ACCCTCACAG AGCCCTGGGT ATGCTTGACC 1980  
ATTGAGGGCC AGGAAGGTTG TCTCCTGGAC ACTGGTGCGG TCTTCTTAGT CTTACTCTTC 2040  
TGTCCCGGAC AACTGTCCTC CAGATCTGTC ACTATTCTGA GGGGGTCCNT AAGACGGGCA 2100  
GTCACTAGAT ACTTTTTCCC AGCCACTAAG TTATGAACTG GGGAGCTTTA TTCTTTTCAC 2160  
ATGCTTTTCT AATTATGCTT GAAAGCCCCA CTACCTTGT AGGGAGAGAC ATTCTAGCAA 2220  
AAGCAGGGGC CATTATACAC CTGAACATAG GAGAAGGAAC ACCCGTTTGT TGTNCCCCTG 2280  
CTTGAGGAAG GAATTAATCC TGAAGTCTGG GCAACAGAAG GACAATATGG ACGAGCCAAA 2340  
GAATGCCCGT CCTGTTCAAG TTAACTAAA GG 2372

(2) INFORMATION FOR SEQ ID NO: 11:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7582 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: mRNA (as DNA)

(iii) HYPOTHETICAL: NO



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CAACAATCGG	GATATAAACC	CAGGCATTCTG	AGCTGGCAAC	AGCAGCCCCC	CTTTGGGTCC	60
CTTCCCTTTG	TATGGGAGCT	GTTTTTCATGC	TATTTCACTC	TATTAAATCT	TGCAACTGCA	120
CTCTTCTGGT	CCATGTTTCT	TACGGCTCGA	GCTGAGCTTT	TGCTCACCCT	CCACCACTGC	180
TGTTTGCCAC	CACCGCANAC	CTGCCGCTGA	CTCCCATCCC	TCTGGATCCT	GCAGGGTGTC	240
CGCTGTGCTC	CTGATCCAGC	GARGCGCCCA	TTGCCGCTCC	CAATTGGGCT	AAAGGCTTGC	300
CATTGTNCCT	GCACGGCTAA	GTGCCTGGGT	TTGTTCTAAT	TGAGCTGAAC	ACTANTCACT	360
GGGTTCCATG	GTTCTCTTCT	GTGACCCACC	GCTTCTAATA	KAACATAAAC	ACTTACCACA	420
TGGCCCAAGA	TTCCATTCCCT	TGGAATCCGT	GAGGSCAACG	AACTCCAGGT	CAGAGAATAC	480
GARGCTTGCC	ACCATCTTGG	AAGCGGCCTG	CTACCRCTCT	GGAAGTGGTT	CACCACCATC	540
TTGGGAGCTC	TGTGAGCAAG	GACCCCCCGG	TRACATTTTG	GCRACCAMSR	ACGGACATCC	600
MAAGTGATGG	GAAACGTTCC	CCGCAAGACA	AAAACGCCCC	TAAGACGTAT	TCTGGARAAT	660
TGGGAMCAAT	TTGACCCTCA	GACACTAAGA	AAGAAACGAC	TTATATTCTT	CTGCAGTGCC	720
GCCTGGCACT	CCTGAGGGAA	GTATAAATTA	TAACACCATC	TTACAGCTAG	ACYTCTTTTG	780
TAGAAAAGGC	AAATGGAGTG	AAGTGCCATA	AGTACAAACT	TTCTTTTCAT	TAAGAGACAA	840
CTCACAATTA	TGTAAAAAGT	GTGATTTATG	CCCTACAGGA	AGCCTTCAGA	GTCTACCTCC	900
CTATCCCAGC	ATCCCCGACT	CCTTCCCCAM	YTAATAAGGA	CCCCCCTTCA	ACCCAAATGG	960
TCCAAAAGGA	GATAGACAAA	AGGGTAAACA	GTGAACCAAA	GAGTGCCAAT	ATCCCCCAAT	1020
TATGACCCTT	CCCAAGCAGT	GGGAGGAAGA	GAATTCGGCC	CAGCCAGAGT	GCATGTGCTT	1080
TTTYTCTCC	CAGACTTAAA	GCAAATAAAA	ACAGACTTAG	GTAAATTCTC	AGATAAYCCT	1140
GATGGCTATA	TTGRTGTTTT	ACAAGGGTTA	GGACAATTCT	TTGATCTGAC	ATGGAGAGAT	1200
ATATATGTCA	CTGCTAAATC	AGACACTAAC	CCCAAATGAG	AGAAGTGCCA	CCATAACTGC	1260
AGCCTGAGRG	TTTGGCGATC	TCTGCTATCT	CAGTCAGGTC	AATGGATANG	GATGACAACA	1320
GAAGGAAAGA	NAATGATTCC	CCACAGGCCA	GCARGCAGTT	CCCAGTCTAS	ACCCTCATTG	1380
GGGACACAGA	AATCAGTAAC	ATGGGAGATT	GGTGCTGCAG	ACATTTGCTA	ACTTGTGTGC	1440
TASAAGGACT	AAGGAAAACCT	ASGAAGAAAR	TCTAYGAATT	ACTCAATGAT	GTCCACCATA	1500
ACACAGGGGA	AGGGAAGAAA	ATCCTACTGC	CTTTCTGGAG	AGACTAAGGG	AGGCATTGAG	1560
GAAGCGTGCC	TCTCTGTAC	CTGACTCTTC	TGAAGGCCAA	CTAATCTTAA	AGCGTAAGTT	1620
TATCACTCAG	TCAGCTGCAG	ACATTAGAAA	AAACTTCAAA	AGTCTGCCCT	AGGCCCCGAG	1680
CAAAACTTAG	AAACCCTATT	GAACTTGGCA	ACYTCGGTTT	TTTATAATAG	AGATCAGGAG	1740
GAGCAGGCGG	AACAGGACAA	ACGGGATTAA	AAAAAAGGCC	ACCGCTTTAG	TCATGACCCT	1800
CAGGCAAGTG	GACTTTGGAG	GCTCTGGAAG	AGGGAAAAGC	TGGGCAAATT	GAATGCCTAA	1860
TAGGGCTTGC	TTCCAGTGCG	GTCTACAAGG	ACACTTTAAA	AAAGATTGTC	CAAGTAGAAG	1920
TAAGCCGCCC	CTTCGTCCAT	GCCCCTTATT	TCAAGGGAAT	CACTGCAAGG	CCCACTGCCC	1980
CAGGGGACAA	AGGTCTTTTG	AGTCAGAAGC	CACTAACCCAG	ATGATCCAGC	AGCAGGACTG	2040
AGGGTGCTTG	GGGCAAGCGC	CATCCCATGC	CATCACCTTC	ACAGAGCCCT	GGGTATGCTT	2100

GACCATTGAG	GGCCAGGAAG	GTTGTCTCCT	GGACACTGGT	GCGGTCTTCT	TAGTCTTACT	2160
CTTCTGTCCC	GGACAACTGT	CCTCCAGATC	TGTCACTATT	CTGAGGGGGT	CCNTAAGACG	2220
GGCAG1CACT	AGATACTTTY	TCCCAGCCAC	TAAGTTATGA	ACTGGGGAGC	TTTATTCTTT	2280
TCACATGCTT	TTCTAATTAT	GCTTGAAAGC	CCCACTACCT	TGTTAGGGAG	AGACATTCTA	2340
GCAAAAGCAG	GGGCCATTAT	ACACCTGAAC	ATAGGAGAAG	GAACACCCGT	TTGTTGTNCC	2400
CCTGCTTGAG	GAAGGAATTA	ATCCTGAAGT	CTGGGCAACA	GAAGGACAAT	ATGGACGAGC	2460
CAAAGAATGC	CCGTCCTGTT	CAAGTTAAAC	TAAAGGATTC	CACCTCCTTT	CCCTACCAAA	2520
GGCAGTACCC	CCTCAGACCC	AAGGCCCAAC	AAGGATTCCA	AAAGATTGTT	AAGGACTTAA	2580
AAGCCCAAGG	CTTAGTAAAA	CCATGCATAA	CTCCCTGCAG	TAATTCCGTA	GTGGATTGAG	2640
GAGGCACAGA	AACCCAGTGG	ACAGTGGAGG	GTTAGTGCAA	GATCTCAGGA	TTATCAATGG	2700
AGGCCGTTGT	CCTTTTATAC	CCAGCTGTAC	CTAGCCCTTA	TACTGTGMYT	TCCCAAATAC	2760
CAGAGGAAGC	AGAGTGGTTT	ACASTCCTGG	ACCTTMAGGA	TGCCTTCTTC	TGCATCCCTG	2820
TACATCCTGA	CTCTCAATTC	TTGTTTGCTT	TTGAAGATAC	TTCAAACCCA	RCATCTCAAC	2880
TCACCTGGAC	TRTTTTACCC	CAAGGGTTCA	GGGATAGYCC	CCATCTATTT	GGCCAGGCAT	2940
TAGCCCAAGA	CTTGAGYCAR	TYMTCATACC	TGGACACTCT	TGTCCTTCRG	TAKGTGGATG	3000
ATTTACTTTT	RGYGCCYRT	TCAGAAACCT	TGTGCCATCA	AGCCACCCAA	GCRCTCTTMA	3060
ATTTCTCGC	YACCTGTGGC	TACAWGGTTT	CCAAACSARA	RGCTCARCTC	TGCTCACAGC	3120
AGGTAAATA	CTTAGGRCTA	ARATTATCCA	AAGGCACCAR	GGCCCTCAGT	GAGGAAYRYA	3180
TCCAGCCTAT	ACTGGCTTAT	CCTCATCYCA	AAACCCTAAA	GCAACTAAGR	GRRTTCCTTG	3240
GCRTAAYAGG	YTTCTGCCGA	AWATGGATTG	CCCAGGTWTG	GCRAAATAGC	CAGGYCATTG	3300
WATACASTAA	TTAAGGAAAC	TCAGAAAGCC	AATACCCATT	TARTAAGATG	GAYAMCTGAA	3360
GYMRAAGTGG	CTTTCCAGGC	CCCTAAAGAA	GGCCTTAAAC	CCAAGYCCCA	GTGTTAAGYT	3420
TGCCAACRGG	GCAAGACTTT	TSTTYATAYR	TCACAGAAAA	AAACAGRAAY	AGCTCTRGA	3480
GTCCTTACAC	AGRTCCRAGG	GAYGAGCTTG	CAACCYRTGG	CRYACCTGAS	TAAGGAAAYT	3540
GATGTAGTGG	CAAAGGGTTG	RCYTCATTGT	TTAYGGGTAG	TGGTGGCAGT	AGCAGTYKTA	3600
GTATCTGAAG	CAGTTAAAAT	AATACAGGGR	AGAGATCTTA	CTGTGTGGAC	ATCTCATGAK	3660
GTGAAYRGCA	TACTCACTGC	TAAAGGAGAC	TTGTGGCTGT	CAGACAACYG	TTTACTTAAA	3720
TRTCAGGCTC	TATTACTTGA	ARGGCCAGTG	CTGCRACTGT	GCACTTGTGC	AACTCTTAAC	3780
CCAGYCNCAT	TTCTTCCAGA	CAATGAAGAA	AAGATARAAY	ATAACTGTCA	ACAARTAAAT	3840
TCTCAAACCT	ATGCCACTCG	AGGGGACCTT	YTAGARGTTC	CYTTGACTGA	TCCYGACCTT	3900
CAACTTGTAT	ACTGATGGAA	GTTCTTTTGT	AGAAAAAGGA	CTTCGAAAAG	YGGGGTATGC	3960
AGTGGTCAGT	GATAATGGAA	TAYTTGAAAG	TAATCCCCTC	ACTCCAGGAA	CTAGTGCTYA	4020
GCTRGACAGAA	CTAATAGCCY	TCAYTKGGGC	ACTAGAATTA	GGAGAAGRAA	AAAGGGYAAA	4080
TATATATACA	GACTCTRART	ATGCTYACCT	AGTCNTCCAT	GCCCATGMRG	CAATATGSAR	4140
AGAAAGGGAA	TTCCTAACTT	CYGAGRGAAC	ACCTATCAMA	CATCAGGAAG	CCATTAGGAR	4200
ATTATTAYTG	GCWGTACAGA	AACCTARAGA	GGTGMAGTC	TTACTACTGCY	GGGGTCATCA	4260

NAAAGGAAAG RAAAGGGAAA TASAAGRGAA YTGCCAAGCA KATATTGAAG CMAAAAGAGC 4320  
 TGCAAGGCAG GACCCTCCAT TAGAAATGCT TATTAACTT CCCTTAGTAT AGGGTAATCC 4380  
 CTTCCGGGAA ACCAAGCCCC AGTACTCAGC AGGAGAAACA GAATGGGGAA CCTCACGAGG 4440  
 CAGTTTTCTC CCCTCGGGAC GGTTAGCCAC TGAAGAAGGG AAAATACTTT TGCCTGCAAC 4500  
 TATCCAATGG AAATTACTTA AAACCCTTCA TCAAACCTTT CACTTAGGCA TCGATAGCAC 4560  
 CCATCARATG GCCAAATCAT TATTTACTGG ACCAGGCCTT TTCAAACTA TCAAGCARAT 4620  
 AKTCAGGGCC TGTGAAKTGT GCCARARAAA TAATCCCCTG CCTYATCGCC AAGCTCCTTC 4680  
 AGCARAACAA ARAACAGGCC ATTACCCTGR ARAARACTGG CAACTGATTT TACCCACAAG 4740  
 CCCAAACCTC AGGGATTTC A GTATCTACTA GTCTGGGTAR ATACTTTCAC GGGTTGGGCA 4800  
 RAGGCCTTCC CCTGTAGGAC AGAAAAGGCC CAAGAGGTAA TAAAGGCACT AGTTCATGAA 4860  
 ATAATTCCCA GATTGCGACT TCCCGAGGC TTACAGAGTG ACAATAGCCC TGCTTTCCAG 4920  
 GCCACAGTAA CCCAGGGAGT ATCCCAGGCG TTAGGTATAC GATATCACTT AACTGCGCC 4980  
 TGAAGGCCAC AGTCCTCAGG GAAGGTCGAG AAAATGAATG AAAYACTCAA AGGACATCTA 5040  
 AAAAAGCAAA CCCAGGAAAC CCACCTCACA TGGCCTGYTC TGTTCCTAT AGCCTTAAAA 5100  
 AGAATCTGCA ACTTTCCCA AAAAGCAGGA CTTAGCCCAT ACGAAATGCT GTATGGAAGG 5160  
 CCCTTCATAA CCAATGACCT TGTGCTTGAC CCAAGACAGC CAACTTAGTT GCAGACATCA 5220  
 CCTCCTTAGC CAAATATCAA CAAGTTCTTA AAACATTACA AGGAACCTAT CCCTGAGAAG 5280  
 AGGGAAAAGA ACTATTCCAC CCWWTGACA TGGTATTAGT CAAGTCCCTT CYCTCTAATT 5340  
 CCCCATCCCT AGATACATCC TGGGAAGGAC CCTACCCAGT CATTTTATYT ACCCCAACTG 5400  
 CGGTTAAAGT GGCTGGAGTG GAGTCTTGGA TACATCACAC TTGAGTCAA TCCTGGATAC 5460  
 TGCCAAAGGA ACCTGAAAAT CCAGGAGACA ACGCTAGCTA TTCCTGTGAA CCTCTAGAGG 5520  
 ATTTGCGCCT GCTCTTCAAA CAACAACCAG GAGGAAAGTA ACTAAAATCA TAAATCCCCC 5580  
 ATGGSCTCC CTTATCATAT TTTTCTCTKT ASTGTTSTTT YACCCTSTTT CACTCTCACT 5640  
 GCACCCCTC CATGCCGCTG TATGACCAGT AGCTCCCCTY ACCMAGAGTT TCTATGGAGA 5700  
 ATGCAGCGTC CCGGAAATAT TGATGCCCCA TCGTATAGGAG TCTTTSTAAG GGAACCCCC 5760  
 ACCTTCACTG CCCACACCCA TATGCCCCGC AACTGCTATC ACTCTGCCAC TCTTTGCATG 5820  
 CATGCAAATA CTCATTATTG GACAGGAAAA ATGATTAATC CTAGTTGTCC TGCAGGACTT 5880  
 GGAGTCACTG TCTGTTGGAC TTACTTCACC CAACTGGTA TGTCTGATGG GGGTGGAGTT 5940  
 CAAGATCAGG CAAGAGAAAA ACATGTAAAA GAAGTAATCT CCCAACTCAC CSGGGTACAT 6000  
 GGCACCTCTA GCCCCTACAA AGGACTAGAT CTCTCAAAAC TACATGAAAC CCTCCGTACC 6060  
 CATACTCGCC TGGTAAGCCT ATTTAATACC ACCCTCACTG GGCTCCATGA GGTCTCGGCC 6120  
 CAAAACCCTA CTAAGTGTG GATATGCCTC CCCCTGAACT TCARGCCATA TGTTTCAATC 6180  
 CCTGTACCTG AACAATGGAA CAACTTCAGC ACAGAAATAA ACACCACTTC CGTTTTAGTA 6240  
 GGACCTCTTG TTTCCAATST GGAAATAACC CATACCTCAA ACCTCACCTG TGTAATAATT 6300  
 AGCAATACTA CATACACAAC CAACTCCCAA TGCATCAGGT GGGTAACCTC TCCCACACAA 6360  
 ATAGTCTGCC TACCCTCAGG AATATTTTTT GTCTGTGGTA CCTCAGCCTA TCGTTGTTTG 6420

09446024 12169

```

AATGGCTCTT CAGAATCTAT GTGCTTCCTC TCATTCTTAG TGCCCCCYAT GRCCATCTAC 6480
ACTGAACAAG ATTTATACAG TTATGTCATA TCTAAGCCCC GCAACAAAAG AGTACCCATT 6540
CTTCCTTTTG TTATAGGAGC AGGACTGCTA GGTGCACTAG GTACTGGCAT TGGCGGTATC 6600
ACAACCTCTA CTCAGTTCTA CTACAAACTA TCTCAAGAAC TAAATGGGGA CATGGAACGG 6660
GTCGCCGACT CCCTGGTCAC CTTGCAAGAT CAACTTAACT CCCTAGCAGC AGTAGTCCTT 6720
CRAAATCGAA GAGCTTTAGA CTYGCTAACC GCTGARAGAG GGGGAACCTG TTTATTTTAA 6780
GGGGAAGAAT GCTGTTATTA TGTTAATCAA TCCGGAATCG TCACTGAGAA AGTTRAAGAA 6840
ATTCSAGATC GAATACAACG TAKAGCAGAR GAGCTTCGAA AACTGGGACC CTGGGGCCTC 6900
CTCAGCCRAT GGATGCCCTG GATTCTCCCC TTCTTAGGAC CTCTAGCAGC TATAATATTG 6960
CTACTCCTCT TTGGACCTG TATCTTTRAC CTCCTTGTTA ACTTTGTCTC TTCCAGAATC 7020
GAAGCTGTRA AACTACAAAT GGAGCCCAAG ATGCAGTCCA AGACTAAGAT CTACCGCAGA 7080
CCCCTGGACC GGCCTGYTAG CCCACGATCT GATGTTAATG ACATCAAAGG CACCCCTCCT 7140
GAGGAAATCT CAGCTGCACA ACCTCTACTA CGCCCCAATT CAGCAGGAAG CAGTTAGAGC 7200
GGTSGTCGGC CAACCTCCCC AACAGCACTT AGGTTTTCTT GTTGAGATGG GGGACTGAGA 7260
GACAGGACTA GCTGGATTTC CTAGGCTGAY TAAGAATCCY TAAGCCTAGS TGGGAAGGTG 7320
ACCACATCCA CCTTTAAACA CGGGGCTTGC AACTTAGYTC ACACCTGACC AATCAGAGAG 7380
CTCACTAAAA TGCTAATTAG GCAAAGACAG GAGGTAAAGA AATAGCCAAT CATYTATTGC 7440
MTGAGAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC CCAAGTYTTC GAGCCGGCAA 7500
CGGCAACCCC CTTTGGGTCC CCTCCCTTTG TATGGGAGCT CTGTTTTTCAT GCTATTCAC 7560
TCTATTAAAT CTGCACTG CR 7582

```

(2) INFORMATION FOR SEQ ID NO: 12:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2563 base pairs
  - (B) TYPE: nucleotide
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

ACTGCACTCT TCTGGTCCAT GTTGTGTACG GCTCGAGCTG AGCTTTTGCT CGCCATCCAC 60  
CACTGCTGTT TGCCACCGTT GCAGACCCAC TGCTGACTTC CATCCCTCTG GATCTGGCAG 120  
GGTGTCTGCT GTGCTCCTGA TCCAGCGAGG GGCCCATTC CACTCCCAAT CGGGCTAAAG 180  
GCTTGCCATT GTTCCTGCAT GGCTAAGTGC CCAGGTTTCAT CCTAATTGAG CTGAACACTA 240  
GTCACTGGGT TCCACAGTTC TCTTCCATGA ACCACGGCTT TTAATAGAGC TATAACACTC 300  
ATCGCAAGGC CCAAGATTCC ATTCCTTGGA ATCTGTGAGG CCAAGAACCC TAGGTCAGAG 360  
AACACGAGGC TTGCCACCAT CTTGGAAGCA GCCTGCCACC ATCTGGGAAG CGGCCTGCCA 420  
CCATCTTGGA AGCCGCCCCG CACCATCTTG GGAGCTCTGG GAGCAAGGAC CTCCCCGCAA 480  
CCCAGTAACA TTAGCGACC ACGAAGGGAC CTCAAAGCG GTAATATTGG ACCACTTTCA 540  
CTTGCTATTC TGCCTATCC TTCCTTAGAA TTGGAGGAAA ATACCGGACA CCTGTCGGCC 600  
GGTTAAAAAC GATTAGCGTG GCCTCCGGAC TTAACAATCA GGTGTGAGGC TATCTGGGGA 660  
AGGGCTTTCT AACAAACCCC AACCRITCTG GGTGGGAAT GTTGGTCTGC CTGGAGCCAG 720  
CTTCCACTTT CAATTTTCCT GGGGAAGCCA AGGGCCGACT AGAGGCAGAA AGCTGTTGTC 780  
CCAAATTCCC GGCAGTAGCC GGTGAGATC ATGGCGCAGC CAGAAGTCTT TACTCCACAG 840  
TCACCCATGC ATGCGCCCCT ATCTTTCCTT CTGACCCATA CCTCCTGGGT CCTAACCATG 900  
ACTTTCTTAA AAGGGTAGCC CCAAATTCT CTTACCTCT GAATCTACTT CCTCTGATCC 960  
CTGCCTCCTA GGTGCTAATG GTTCAGACTT TCATTTCCTC TAGCAAGTTG TATYTCCAAA 1020  
GGGATATAAG GAAGCTCTAC ACTGTATCCT TAGGCATCTA GGCTCTAAAC CCAGGGAGTC 1080

TTGTCCCTGA TGTCCCAACC GATTTAGGTA TATAGTTCTC GACATGGGCA GTTATGTGGG 1140  
ACCCATTCCC CACCACCCTT GCCAGGGCCC CAAGTTTGTA AATGGCTAAG AGAGGAAAGT 1200  
GAGAGAGAGA GAGACAGAGT GAGACACAGA GAGAGGGAGA GACAGAGAGA GAGACAGAGA 1260  
GGAGAGAGAC ACAGAGAGGG GAGAGACACA GAGAGGAGAA GGGGGCAGAG AGACCAAGAG 1320  
GGAGTCYMAG AGAGAGAGAA AGAAGAAGAA ATAGTAGAAA AAAAAGTGTG CCCTATTCCT 1380  
TTAAAAGCCA GGGTAAATTT AAAAACCTA TACTTGATAA TTGAAGGTCT TCTCCATGAC 1440  
CCTGTAACAC TCTAATACTA CTTGTTCTC AGTGTAACA AGGGTGTTAG CCTGAAAACA 1500  
CTGAGACCGC TGACACCCAT AGCTTTCCTA TAAAAATCC TTAACCCAGT AACCCGAGA 1560  
TGGCCCGCAT GCATTCAATC TGTAGTGGCA ACTGCTTTGC TAACAAGAAT AAAGTGGAAA 1620  
AGTAACTTTT AGAGGAAACC TCATTGTGAG CACACCTCAC CAGTTCAGAA TTATTCTAAG 1680  
TCAAAAAGC AAAAAGGTAG CTTACTAACT CAAAATCTT AAAGTATGGG GTTATTTTGT 1740  
TAGAAAAGG TAATTTAACA CTAATCACTG ATAATTCCTT TAACCCAGAA GATTTCCTAA 1800  
CAGGAGATTT AAATCTTAAT TACCATACAA AGGTCTGACC AGACCTAGGA GGAATCCCT 1860  
TCAGTACAGG ATGATAGATG GTTCCTCCCA GGTGAATGAA AAAAAATCA CAATGGGTAT 1920  
TCAGTAATTG ATAGGGAGAC TCTTGTGGAA GCAGAGTTAG AAAAAGTACC TAATAATTGG 1980  
TCTCCCCAAA CCTGCGAGCT GTTTGCACTC AGCCAAGCCT TAAAGTACTT CTAGAATCAA 2040  
AAAGATTATC TCAATCCTGA CTCAAAAGGT TACCTACACC CTCTGTGAAA CGAATTTACT 2100  
TAAGAACTGT TTATGGGACT GCATCTTGAT GGGGCAGCTG GGTGTGTCATG AAATACTCAG 2160

66922"42094450

GAATGCAGCC TAGCTCTAGG ACTCACCCCT GAGCACAAAG GCAATGTTGG GCATGCTGGT 2220  
 AAAGGACCAC TAGAATCCAG CAGTCCGAAC CCTTTCTTTG GGTTAAGAAA GGCGGGAAAA 2280  
 CAGGCGCAGG ACTGCTACAT TGGTAAGCGT AACTAATCCA ATAAGCAGAG GTCCATGGGT 2340  
 GGTGACACAC TCTGGAAAGC AATAAGCATT AGRACCATAG AGGACGCTCT ACGACTAATG 2400  
 CTCGTCGGAA AATGACTAGA GGTGCTGGCA TCCCTATGTT CTTTTTTCAG ATGGGAAATG 2460  
 TTCCCCCTCA AGGCAAAAAC ACCCCTAAGA TGTATTCTGG ACAATTGGGA CCAATTTGAC 2520  
 CCTCAGACTC TAAGAAAGAA ACGACTTATA TTCTTCTGCA GTG 2563

(2) INFORMATION FOR SEQ ID NO: 13:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2585 base pairs  
 (B) TYPE: nucleotide  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: DNA  
 (iii) HYPOTHETICAL: NO

- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCAGGGATAG CCCCCATCTA TTTGGCCAGG TATTAGCCCA AGACTTGAGC CACTTCTCAT 60  
 ACTTGACAC TCTTGTCTT TGGTATGTGG ATGATCTACT TTAGCCACC TGTTAGAAA 120  
 CCTTGTGCCA TCAAGCCAAC CAAGTGCTCT TAACTTCCT CGCCACCTGT GGCTACAAGG 180  
 TTTCAAACC AGAGGCTCAG CTCTGCTTAC AGCAGGTAA ATACTTAGGG CTAAAATTAT 240

669721"42094460

CCAAAGGCAC CAGGGCCCTC AGTGAGGAAC GTATCCAGCC TATACTGGCT TATCCTCATC 300

CCAAAACCCT GAAGCAATTA AGAGGGTTCC TTGGCATAAA AGGCTGCTGT TGAATATGGA 360

TTCCCAGGTA CAATGAAATA GCCAGGCCAT TATACACACT AATTACGGGA ACTCAGAAAG 420

CCAATACCCA TTTAGTAGAA TGGACACCTG AAGCAGAAGC GGCTTTCCAG GCCCTAAAGA 480

AGGCCCTAAT CCAAGCCCCA GTGTTAAGCT TGCCAATGGA GCAAGACTTT TCTTTATATG 540

TCACAGAAAA AAAAACAGGA ATAGCTCTAG AAGTCCTTAC ACAGGTCCGA GGGACCAGCT 600

TACAACACAT GGCATACCTG AGTAAGGAAA CTGATGTAGT GGCAAAGGGT TGGACTCATT 660

GTTTACAGGT AGTGGCAGCA GTAGCAGTCT TAGCATCTGA AGCAGTTAAA ATGATACAGG 720

GAAGANATCT TACTGTGTGG ACATCTCATG ATGTGAACGG CATACTCACT GCTAAAGGAG 780

ACTGTGGCTG TCAGACAACC ATTTGCTTAA ATATCAGGCT CTATCACTTG AANGGCCAGT 840

GCTGCCACTG TGCACTTGTG CAACTCTTAA CCCACCCACA TTTCTTCCAG ACAATGAAGA 900

AAAGATAGAA CATAACTGTC AACAAAGTGAT TGTTCAAACC TACACCGCTC GAAGGGACCT 960

TCTAGAGGTT CCCTTGACTG ATCCTGAGCT CAACTTCTAT ACTGATGGAA GTTCCTTTTG 1020

TAGAAAAAGG ACTTCGAAAG GCGGGTATGC AGTGGCCAGT GATAATGGAA TACTTGAAAG 1080

TAATCCCTTC ACTCCAGAAA CTAGCATTCA GCTGGCAGAA TTAATAGCCT TCACTTGGGC 1140

ATTAGAACAC AGGAGAAGGA AAAGGAGTAA ATATATATAC AGACTCCAAG TATGCTTACT 1200

TAGTCCTCCA TGCCCATGCA GCAATATAGA GAGAAAGCGA ATTCCTAACT TCTGAGGGAA 1260

CACCTATCAA ACATCAGGAA GCCATTAGGA GATTATTACT GGCTGTACAG AAACCTAGAG 1320

0944604-1269



GTGGCAGTCT TACATGGCCG AGATCATCAG AAAGGAAAAG AAAGGGAAAT AGAAGGGAAC 1380  
TGCCAAGTGG ATATTGAAGC CAAAAGAGCT GCAAGGCGGG ACCCTCCATT AGAAATGCTT 1440  
ATAGAAGGAC CCCTAGTACA GGGCAATCCC CTTCAGGAAA CCAAGCCCCA ATACTCAGCA 1500  
GAAGAAATGG AATGGGGAAC CTCATGAGGA CATAGTTTCC TCCCCTCAGG ATGGCTAGCC 1560  
ACCAAAGAAG GAAAAATACT TTTGCCTGCA GCTAACCAAT GGAAATTACT TAAAACCCTT 1620  
CACCAAACCT TTCGCTTAGG CATTGATAGC ACCCATCAGA TGGCTAAATC ATTATTTACT 1680  
AGACCACACC TTTTCAAAAC TATCAAGCAG ACAGTTAGGG CCTGTGAAGT GTGCCAAAGA 1740  
AATAATCCCC TGCCTTATCG CCAAACCTCT TCAGGAGAAA AAAGAACAGG CCATTACCCA 1800  
GGAGAAGAGT GGCAACTAGA TTTTACCCAC ATGCCCAAAT CTCAGGGATT TCAGTATCTA 1860  
CTAGTCTGGG TAGATACTTT CACTGGTTGG GCGGAGGCCT TCCCTTGTAG GACAGAACAG 1920  
GCCCATGAGG TAATAAAGGC ACTAATTCAT GAAATAATC CCAGATTTGG ATTTCCCCAA 1980  
GGCTTACAGA GTGATAACGG CCCCACTTTC AAGGCTACAG TAACCCAGGG AGTATCCCAG 2040  
ACATTAGACA TACAATATCA CTTACACTGA GCCCGGAGGC CACAATCCTC AGGAAAGTTG 2100  
AGAAATGAA TGAAACGCTC AAATGACATC TAAAAAGCT AACCTAAGAA ACCCACCTCT 2160  
CATGGTTTGC TCTGTTGCCT ATAGCCTTAG TAAGAATCCG AAACCTCTCC CAAAAGCGG 2220  
GACTCAGCCC ATACGAAATG CTGTATGGAC GGCCCTTCCT AACCAATGAC CTTGTGCTTG 2280  
ACCTAGAGAT GGCCAACTTA GTTGCAGATA TCCCTCCTTA GCCAAATATC AACAGTTCT 2340  
TAAAACGTCA CAGGGAACCT GTCCCTGAGA GGAGGGAAAG GAATTATTCC AACCTGGTGA 2400

CATGGTATTA GTGAAGTCCC TTCCCTCCAA CTCCCCATCC CCTGGATACA TCCTGGGAAG 2460  
GACCCTACTC AGTCATTTTA TCTATCCCAA CCGCGGTAA AATGGCTGGA GTAGAATCTT 2520  
GGATACATCA CATTGAGTC AAACCCTAGA TACTGCCACA AGGAACCTGA AAATCCAGGA 2580  
GACAA 2585

(2) INFORMATION FOR SEQ ID NO: 14:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2575 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: NO

- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGGATAGCCC CCATCTATTT GGCCAGGCAT TAGCCCAAGA CTTGAAGCCA ATTCTCATAC 60  
CTGGACACTC TTCTCCTTTG GTATGTGGAT GATTACTTT TAGCTTCCTG TTCAGAAACC 120  
TTGTGCCATC AAGCCACCCA AGCACTCTTA AATTCCTCG CTACCTGTGG CTACAAGGTT 180  
TCCAAACCAA AGACCCAGCT CTGCTCACAG CAGGTAAAT ACTTGGGGCT AAAATTATCC 240  
AAAGGCACCA GGGCCCTCAG TGAGGAACGT ATCAAGCCTA TACTGGCTTA TCCTCATCCC 300  
CAATCCTAA AGCAACTAAG AGAGTTCCTT AGCATACAG GTTCTGCTG AATATGGATT 360

CCCAGGTATG GCAAAATAGC CAGACCATTA TATACGCTAA TTAAGGAAAC TCAGAAAGCC 420  
 AATACCCATT TAGTAAGATG GATACCTGAA GCAGAAGCAG CTTTCCAGGC CCTAAAGAGG 480  
 GCCCTAACCC AAGCCCCAGT GTTAAGCTTG CCAACAGGGC AAGACTTTAC TTCGTATGTC 540  
 ACAGAAAAAA CAGGAAATAG CTCTAGGAGT CCTTACACAA GTCTGAGGGA TGAGCTTGCA 600  
 ACCCATGGCA TACCTGAGTA AGGAAATTGA TGTAGTGGCA AAGGGTTGGC CTCATTGTTT 660  
 ATGGGTTAGTG GCGGCAGTAG CAGTCTTAGC ATCTGAAGCA GTTAAAATGA TACAGGGAAG 720  
 AGATCTTACT GTGTGGACAT CTCATGATGT GAATGGCATA CTCACTGCTA AAGGAGACTT 780  
 GTGGCTGTCA GACAACCATT TACTTAAATA TCAGGCTGTA TTACTTGAAG GGCCAGTGCA 840  
 GCAACTGCGC AGTTGTGCAG CTCTTAACCC AGCCACATTT CTTCCAGACA ATGAAGATAG 900  
 AACATAACTG CCAACAAGTA ATTTCTCAA CCTAGGCCGC TCGAGGGAAC CTTTTAGAGG 960  
 TTCCCTTAAC TGATCCCGAC CTCAACTTGT ATACTGATGG AAGTTCCTTT GTAGAAAAAG 1020  
 GACTTTGAAA AGTGGGGTAT GCAGTGCTCA GTGATAATGG AATACTTGAA AATAATCCCT 1080  
 TCATTCCAGG AACCAGCGTT CAGCTGGCAG AATTAATAGC CCTCACTCGG GCATTAGAAT 1140  
 TAGGAGAAGG AAAAAGGGTA AATACACATA CAGATTCTAA GTATGTTTAC TTAGTCCTCC 1200  
 GTGCCCACGC AGCAATATGG AGAGAAAGGG AATGCTTAAC TTCTGAGGGA ACACCTATCA 1260  
 AACATCAGGA AGTTATTAGG AGATTATTAT TGGCTATACA GAAACCTAAA GAGGTGGCAG 1320  
 TCTTACTCTG CTGGGGTGGT CAGAAAGAAA AGGAAAGGGA AATAAAAGGG AACTGCCAAG 1380  
 CGGATATTGA AGCCAAAAGA GCCGCAAGGC AGGACCCTCC ATTAGAAATG CTTATAGAAG 1440

66922" 4209460

GACCCCTAGT ATGGGGTAAT CCCCTCCGGG AAACCAAGCC CCAATACTTA GAAAAAGAAA 1500  
TAGAATGGGG AACCTCACGA GGACATAGTT TCCTCCCCTC AGGATGGCTA GCCACCGAAG 1560  
AAGGAAAAAT ACTTTTGCCT GCAGCTAACC AATGGAAATT ACTTAAAACC CTTACACAAA 1620  
CCTTTCACCT AGACATTGAT AGCACCCATC AGATGGCCAA ATCATTATTT ACTGGACCAG 1680  
GCCTTTTCAA AACTATCAAG CAGCTAGTCA GGGCCTGTGA AGTGTGCCGA AGAAATAATC 1740  
CCATGCCTTA TCACCAAGCT CCTTCAGGAG AACAAAGAAC AGGCCATTAC CCAGCAGAAG 1800  
RVTGGCAACT AGATTTTACC CACATGCCCA AATCTCAGGG ATTTCAGTAT CTACTAGTTT 1860  
GGGTAGATAC TTTCACCTGGT TGGGCAGAGA CCTTCCCCTG TAAGACAGAA AAGTCCCAG 1920  
AGGTAATAAA GGCATTAGTT CATGAAATAA TTCCAGATT CAGACTTCCC TGAGGCTTAC 1980  
AGAGTGACAA TGGCCCTGCT TTCAAGGCTA CAGTAACCCA GGAGTATCCC AGGTGTTAGG 2040  
TATACAATAT CACTTACACT GCGCCTGGAG GCAGTCCTCA GGAAGGCCG AGAACTGAA 2100  
TGAAACACTC AAACGACATC TAAAAAAGC TAACCAGGA AAACCACCTC ACATGGCCTG 2160  
CTCTGTTGCC TATAGCCTTA CTAAGAATCC AAAACTCTCC CAAAAAGCA GGACTTAGCC 2220  
CATACGAAAT GCTATATGGA TAGCCCTTCC TAACCAATGA CCTGTGCTT GACTGAGAGA 2280  
GAGCCAACTT AGTTGCAGAC ATCACCTCCT TATCCAAATA TCAACAAGTT CTTAAAACAT 2340  
TACAAGGAGC CTGTCCCCGA GAAGAGGGGA AGGAATATT CCACCCTGGT GACATGGTAT 2400  
TAGTCAAGTC CCTTCCCTCT AATTCTCATT GCCTAGATAT ATCCTGGGAA GGACCCTACC 2460  
CAGTCATTTT ATCTACCCCA ACCGCAGTAA AAGTGCTGG AGTGGAGTCT TGGATACATC 2520  
ACACTCGAGT CAAACCCTGG ATATTACCAA AGGAACCTGA AAATCCAGGA GACAA 2575

66922-4209460

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 783 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

10

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

669727"42094460

TGAGAGACAG	GACTAGCTGG	ATTCCTAGG	CYACTAAGA	ATCCYTAAGC	CTAGSTGGGA	60
AGGTGACCAC	RTCCACCTTT	AAACACGGGG	CTTGCAACTT	AGYTCACACC	TGACCAATCA	120
GAGAGCTCAC	TAAAATGCTA	ATTAGGCAAA	GACAGGAGGT	AAAGAAATAG	CCAATCATYT	180
ATTGCMTGAG	AGCACAGCAG	GAGGGACAAY	RATCGGGATA	TAAACCCARG	YHTTCGAGCY	240
GGCAACRGCA	GMCCCCCTTT	GGGTCCCYTC	CCTTTGTATG	GGAGCTCTGT	TTTCATGCTA	300
TTTCACTCTA	TTAAATCTTG	CARCTGCRCT	CTTCTGGTCC	ATGTTTCTTA	CGGCTYGAGC	360
TGAGCTTTYG	CTCRCCRTCC	ACCACTGCTG	TTTGCCRCCA	CCGCANACCY	GCCGCTGACT	420
CCCATCCCTC	TGGATCMTGC	AGGGTGTCCG	CTGTGCTCCT	GATCCAGCGA	RGCRCCCAT	480
GCCGCTCCCA	ATYGGGCTAA	AGGCTTGCCA	TTGTNCCTGC	AYGGCTAAGT	GCCTGGGTTY	540
RTYCTAATTG	AGCTGAACAC	TANTCACTGG	GTTCCATGGT	TCTCTTCTGT	GACCCACRGC	600
TTCTAATAGA	RCTATAACAC	TYACCRCATG	GCCCAAGRTT	CCATTCCCTG	GAATCCRTRA	660
RGSCAACGAA	CYCCASGTCA	GAGAAYACGA	RGCTTGCCAC	CATCTTGGA	GCGGCCTGCT	720
ACCATCTTGG	AAGTGGTTCA	CCACCATCTT	GGGAGCTCTG	TGAGCAAGGA	CCCCMRGTR	780
ACA						783

15

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

5 TGTCCGCTGT GCTCCTGATC 20

(2) INFORMATION FOR SEQ ID NO: 17:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATGCACTCTG GCTGGGCCAA T 21

(2) INFORMATION FOR SEQ ID NO: 18:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

669727" 42094460

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ACCATTTGAC CCTCAGACAC T

21

5 (2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleotide

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

15 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AACCCTTTGC CACTACATCA ATTT

24

20

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

25 (B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

30

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

35 TCAGGGATAG CCCCCATCTA T

21

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

15 TTGTCTCCTG GATTTTCAGG TT 22

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

30 GGACCCTACC CAGTCATTTT 20

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleotide



(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

5

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

10

ATCAGGAGCA CAGCGGACAC

20

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 22 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

25

GGACATCCAA AGTGATACAT CC

22

(2) INFORMATION FOR SEQ ID NO: 25:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA

669727-4209460

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

5 AATGTATGGC CTGAAGTGCA G

21

(2) INFORMATION FOR SEQ ID NO: 26:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CTTCCCAGGA TGTATCACTT TG

22

(2) INFORMATION FOR SEQ ID NO: 27:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CACTGCAGAA GAATATAAGT CGTT

24

(2) INFORMATION FOR SEQ ID NO: 28:

REPLACEMENT SHEET (RULE 26)

009727 172094460

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

5

- (ii) MOLECULE TYPE: DNA

10

- (iii) HYPOTHETICAL: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GCTTCCAAGA TGGTGGCAAG C

21

15

- (2) INFORMATION FOR SEQ ID NO: 29:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 678 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

- (ii) MOLECULE TYPE: DNA

25

- (iii) HYPOTHETICAL: NO

09445024 121699

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

TCAGGGATAG CCCCATCTA TTTGGCCAGG CATTAGCCCA AGACTTGAGC CAGTTCTCAT 60  
ACCTGGATAT TCTTGTCTT TGGTATGCGG ATGATTACT TTTAGCCGCC CGTTCAGAAA 120  
CCTTGTGCCA TCAAGCCACC CAAGTGCTCT TAAATTCCT CGCCACCTGT GGCTACAAGG 180  
TTTCCAAACC AAAGGCTCAG CTCTGCTCAC AGCAGAAGGC TATTACCCT AAATACTTAG 240  
GGCTGAAATT ATCCAAAGGC ACCAGGGCCC TCAGTGAGGA ATGTATCCAG CCTATACTGG 300  
CTTATCCTTA TCCCAAACC CTAAACAAC TAAGAAGGT CTTGGCATA ATAGGCATAA 360  
CAGGCATAAC AGGTTTCTGC TGAATATGGA TTCCAAGTA CGGCAAATA GCCAGACCAT 420  
TATATACACT AATTAAGGAA ACTCAGAAAG CCAATACCCA TTTAGTAAGA TGGACACCTG 480  
AAGCAGAGGC AGCTTTCCAG GCCGTAAAGA ACACCCTAAC CCAAGCCCCA GTGTTAAGCT 540  
TGCCAGCGGG GCAAGACTTT TCTTTCTGTG TCACAGAAA AATAGGAATA GCTNTAGGAG 600  
TCCTTACACA GGTCCGAGGG ACCAGCTTGC AACCCATGGC ATACCTGAGT AAGGAAATTG 660  
5 ATGTAGTGGC AAAGGGTT 678

(2) INFORMATION FOR SEQ ID NO: 30:

- (i) SEQUENCE CHARACTERISTICS:  
10 (A) LENGTH: 536 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
15 (ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCAATCTCCA TGTTGTATCC CCTTCCCCAA CTAATAAGGA CCCCCCTTTC AACCCAAACA 60  
GTCCAAAAGG ACATAGACAA AGGAGTAAAC AATGAACCAA AGAGTGCCAA TATTCCTGG 120  
TTATGCACCC TCCAAGCGGT GGGAGAAGAA TTCGGCCCAG CCAGAGTGCA TGTACCTTTT 180  
TCTCTCTCAC ACTTGAAGCA AATTAAATA GACCTAGGTA AATTCTCAGA TAGCCCTGAT 240  
GGCTATATTG ATGTTTTACA AGGATTAGGA CAATCCTTTG ATCTGACATG GAGAGATATA 300  
ATATTACTGC TAAATCAGAC GCTAACCTCA AATGAGAGAA GTGCTGCCAT AACTGGAGCC 360  
CGAGAGTTTG GCAATCTCTG GTATCTCAGT CAGGTCAATG ATAGGATGAC AACGGAGGAA 420  
AGAGAACGAT TCCCCACAGG GCAGCAGGCA GTTCCCAGTG TAGCTCCTCA TTGGGACACA 480  
GAATCAGAAC ATGGAGATTG GTCCCGCAGA CATTTAAAGC TTTCCCCGGG TACCGA 536

5 (2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 591 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

15 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CCATGGCCAT CTACACTGAA CAAGATTTAT ACAATCATGT CGTACCTAAG CCCCACAACA 60  
AAAGAGTACC CATTCTTCCT TTTGTTATCA GAGCAGGAGT GCTAGGCAGA CTAGGTACTG 120  
GCATTGGCAG TATCACAACC TCTACTCAGT TCTACTACAA ACTATCTCAA GAAATAAATG 180  
GTGACATGGA ACAGGTCACT GACTCCCTGG TCACCTTGCA AGATCAACTT AACTCCCTAG 240  
CAGCAGTAGT CTTTCAAAT CGAAGAGCTT TAGACTTGCT AACCGCCAAA AGAGGGGGAA 300  
CCTGTTTATT TTTAGGAGAA GAACGCTGTT ATTATGTTAA TCAATCCAGA ATTGTCACTG 360  
AGAAAGTTAA AGAAATTCGA GATCGAATAC AATGTAGAGC AGAGGAGCTT CAAAACACCG 420  
AACGCTGGGG CCTCCTCAGC CAATGGATGC CCTGGGTTCT CCCCTTCTTA GGACCTCTAG 480  
CAGCTCTAAT ATTGTTACTC CTCTTTGGAC CCTGTATCTT TAACCTCCTT GTTAAGTTTG 540  
TCTCTTCCAG AATTGAAGCT GTAAAGCTAC AGATGGTCTT ACAAATCTAG A 591

5 (2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 364 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

15 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CTAACCTGAG GATCCAGCAG CAGGACTGAG GGTGCCCCGGG GCAAGTGCCA GCCCATGCCA 60  
TCACCCTCAG AGCCCCGGGT ATGTTTGACC ATTGAGAGCC AGGAAGTTAA CTGTCTCCTG 120  
GACACTGGCG CAGCCTTCTC AGTCTTACTT TCCTGTCCA GACAATTGTC CTCCAGATCT 180  
GTCACTATCC GAGGGGTCCT AGGACAGCCA GTCACTACAT ACTTCTCTCA GCCACTAAGT 240  
TGTGACTGGG GAACTTTACT CTTTTCACAT GCTTTTCTAA TTATGCCTGA AAGCCCCACT 300  
CCCTTGTTAG GGAGAGACAT TTTAGCAAAA GCAGGGGCCA TTATACACCT GAACAAGCTT 360  
GAAA 364

5 (2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 538 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

15 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Met Gly Leu Pro Tyr His Ile Phe Leu Cys Ser Val Leu Ser Pro Cys  
1 5 10 15

Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser Ser Ser  
20 25 30

Pro His Pro Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn Ile Asp  
35 40 45

Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe Thr Ala  
50 55 60

His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu Cys Met  
65 70 75 80

His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro Ser Cys  
85 90 95

Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr Gln Thr  
100 105 110

Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu Lys His  
115 120 125

Val Lys Glu Val Ile Ser Gln Leu Thr Gly Val His Gly Thr Ser Ser  
130 135 140

Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu Arg Thr  
145 150 155 160



His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly Leu His  
165 170 175

Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu Pro Leu  
180 185 190

Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp Asn Asn  
195 200 205

Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro Leu Val  
210 215 220

Ser Asn Val Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val Lys Phe  
225 230 235 240

Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp Val Thr  
245 250 255

Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe Val Cys  
260 265 270

Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser Met Cys  
275 280 285

Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu Gln Asp  
290 295 300

Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val Pro Ile  
305 310 315 320

Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly Thr Gly  
325 330 335

Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu Ser Gln  
340 345 350

Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val Thr Leu  
355 360 365

Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Arg Asn Arg Arg  
370 375 380

Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu Phe Leu  
385 390 395 400

Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val Thr Glu  
405 410 415

Lys Val Glu Glu Ile Pro Asp Arg Ile Gln Arg Ile Ala Glu Glu Leu  
420 425 430

Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Arg Trp Met Pro Trp Ile  
435 440 445

Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu Leu Phe  
450 455 460

Gly Pro Cys Ile Phe Asp Leu Leu Val Asn Phe Val Ser Ser Arg Ile  
465 470 475 480

Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys Thr Lys  
485 490 495

Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser Asp Val  
500 505 510

Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala Gln Pro  
515 520 525

Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser  
530 535

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 52 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Met Glu Pro Lys Met Gln Ser Lys Thr Lys Ile Tyr Arg Arg Pro Leu  
1 5 10 15

Asp Arg Pro Ala Ser Pro Arg Ser Asp Val Asn Asp Ile Lys Gly Thr  
20 25 30

Pro Pro Glu Glu Ile Ser Ala Ala Gln Pro Leu Leu Arg Pro Asn Ser  
35 40 45

15 Ala Gly Ser Ser  
50

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 48 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Met Leu Met Thr Ser Lys Ala Pro Leu Leu Arg Lys Ser Gln Leu His  
1 5 10 15

Asn Leu Tyr Tyr Ala Pro Ile Gln Gln Glu Ala Val Arg Ala Val Val  
20 25 30

Gly Gln Pro Pro Gln Gln His Leu Gly Phe Pro Val Glu Met Gly Asp  
35 40 45

0944604.12169  
669F2T" 42094460